

DOCUMENT RESUME

ED 292 866

TM 011 273

AUTHOR Jelden, D. L.
TITLE Modes and Types of Computerized Unit Test Item
Feedback and Its Effect on Final Exam Scores.
SPONS AGENCY University of Northern Colorado, Greeley.
PUB DATE 87
NOTE 48p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS College Students; *Computer Assisted Instruction;
Computer Assisted Testing; *Feedback; Higher
Education; Oceanography; Scores; Test Coaching; *Test
Items
IDENTIFIERS Final Examinations; *PHOENIX (Computer System)

ABSTRACT

A study of 696 undergraduates at the University of Northern Colorado was undertaken to determine the effects of computerized unit test item feedback on final examination scores. The study, which employed the PHOENIX computer managed instruction system, included students at all undergraduate levels enrolled in an Oceanography course. To determine the compatibility of the groups, American College Test scores were used in an analysis of covariance to account for initial group ability differences. Mean final grades and final examination scores were also compared. To determine if final examination scores had any relationship to the type and amount of feedback on unit tests, several variations of test feedback on unit test administration were established. Feedback was presented via sequential or random objective mode. Results indicate that: (1) corrective feedback slightly improves students' achievement scores as measured by objective tests administered by computer; (2) the speed of computer-assisted item presentation is essential; (3) immediate feedback on specific test items is more beneficial to students than is delayed feedback (retention of material is enhanced with immediate feedback); (4) low-ability students seem to benefit more than do high-ability students from corrective test feedback; and (5) tests generated by sequential course objective and in random order by item within that objective yield higher scores than do tests created by random objective order. (TJH)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

MODES AND TYPES OF COMPUTERIZED UNIT
TEST ITEM FEEDBACK AND ITS EFFECT
ON FINAL EXAM SCORES

STAFF STUDY

Interdisciplinary Center for Educational Technology

D.L. Jelden, Ed.D.

Professor, Educational Technology

University of Northern Colorado

Greeley, Colorado 80639

1987

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

* This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

* Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

D.L. JELDEN

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

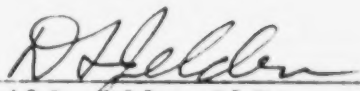
ED 292866

TM 011 273

This Research was Sponsored by
The University of Northern Colorado

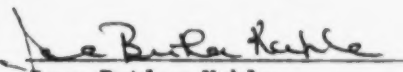
1986-87

Researched By:


David L. Jelden, Ed.D.

Under the Direction of:

Dean, College of Education


Jane Butler-Kahle

Director—ICET


Joseph P. Lounge

Completion of the Project: June 14, 1987

Acknowledgements

This study was undertaken with the support of several university groups. The UNC Computer Center provided the terminals and hardware for the study; the Oceanography classes at UNC 1986-87 conducted by Dr. Bill Hoyt provided the population for the study; the research procedures and testing criterion were validated by Dr. Dale Shaw and Dr. Eugene Koplitz; the data entry and SAS computer packages were completed by Diane Suhr.

Special thanks go to Dr. Bill Hoyt who allowed his students to participate in the study. Finally, I wish to thank the students in OCN 110 who provide the data as a part of their instructional program.

TABLE OF CONTENTS

List of Tables

Chapter 1

Page

Introduction to the problem.....	2
Purpose of the study.....	3
Rationale and assumptions.....	4
Definition of terms.....	5
Limitations of the study.....	6

Chapter 2

Review of Literature.....	7
The nature of test feedback.....	7
Types of feedback and ability level.....	9
Attitudes toward CMI test feedback.....	13

Chapter 3

Methods and Procedures.....	14
-----------------------------	----

Chapter 4

Findings of the study.....	24
Introduction.....	24
Demographic information.....	24
Effects of unit feedback on final exam scores.....	27
Effects of unit feedback on final exam scores.....	31
Mode of presentation.....	31

Chapter 5

Summary of the study, findings, conclusions, implications and	
---	--

recommendations.....	33
Review of literature: nature of test feedback.....	33
Methods and procedures.....	34
Findings of the study.....	38
Conclusions, implications and recommendations.....	42
Bibliography.....	46

LIST OF TABLES

Table 1	Population of the study.....	24
Table 2	Class makeup of the population of the study.....	25
Table 3	Comparison of ACT scores of the population.....	26
Table 4	Final test scores an class grades of groups.....	26
Table 5	Analysis of covariance for final exam scores.....	29

Chapter 1

Introduction

Improving the quality of the teaching-learning process has been a major goal of professional educators. Various methods and techniques for attaining this goal have been developed and tested through the years. Work sheets, study guides, video tapes and films have supplemented the traditional classroom lecture.

Recent observations, however, suggest that the computer holds significant promise for improving the teaching-learning process. The computer has the capability of giving immediate feedback to a students' response to various stimuli. Thus, according to some learning theorists, a student should experience greater gain in learning subject matter than would otherwise be expected utilizing the traditional teacher lecture method. (Kulik, 1983, Bjork, 1984)

Immediate feedback regarding a students' cognitive response to questions related to a subject being studied should improve the students comprehension and speed in mastering cognitive concepts. A computer programmed course incorporates the concept of testing the learner at appropriate intervals as the student progresses through the course material. Responses to test items indicate whether or not the student comprehends the information presented. Immediate test results facilitate the learner to correct misunderstandings or if the responses are correct the effect will create a positive reinforcement of attaining specific knowledge coupled with comprehension.

Whatever the purposes served by testing, very little research exists regarding the instructional value of feedback on test items during test administration, particularly as it relates to computer delivered exams. This research effort investigates some specific questions regarding the instructional value or worth of feedback on test items toward increased subject matter comprehension. In addition, this study attempted to validate assumptions regarding the modes of presentation of test items in objective/item sequence.

Purpose Of The Study

The purpose of this study was to compare the academic performance of college students responding to combinations of test item feedback methods and modes of test item presentation to final examination scores. All examinations and feedback were administered by the PHOENIX computer based education system. More specifically, the study attempted to answer the following questions.

- (1) Will students who receive partial feedback of 'correct' or 'incorrect' and are provided the correct answer on unit test items score higher on final examinations than students who receive no feedback?
- (2) Will students who are presented unit test items in sequential order by course objective score higher on final examinations than students who presented the unit test items in random course objective order?
- (3) What relationship exists between the types and amount of feedback on unit test items presented in random order by course objective to student ACT scores and final examination scores for the course?
- (4) What relationship exists between the types and amount of feedback on unit test items presented in sequential order by course objective to the ACT scores and final examination scores of students in the course.

Hypotheses To Be Tested

- H01 - Students presented unit test items in sequential order by course objective will score significantly higher than students who are presented the unit test items in random course objective order.
- H02 - Students given feedback on unit test items of 'correct' and 'incorrect' as well as the right answer will score significantly higher on final examinations than students who are provided no feedback on unit test items.
- H03 - Students administered unit test items in sequence by course objective and who are given feedback of 'correct' and 'incorrect' as well as the right answer will score significantly higher on final examinations than students given similar information on unit test items but administered in random course objective order.
- H04 - Students administered unit test items in sequential order by course objective and with no feedback will score significantly higher on final examinations than students who are administered similar test items in random course objective order.

Rationale and Assumptions

This study was an attempt to look specifically at the influence of test item feedback on unit study tests to the overall achievement gain of students as measured by an objective final examination. Although related to interactive instruction and computer delivery, very little is known about the specific value of test item feedback during the evaluation mode of course delivery. Traditionally instructors have administered paper and pencil tests, scored them and then discussed in the class the answers for the benefit of student learning. This study dealt with the computer's ability to provide the same type of feedback information to students only the feedback will be presented to the students by the computer during test administration or immediately after all items on the test are taken. If, as suspected, the feedback on unit test items does provide the student with support for decisions made in testing, and the amount and types of feedback increase the

learning of students, then the provision for feedback on computer administered tests should be built into the computer instruction system. The nature of the feedback and mode of presentation for test item administration was the main thrust of this study.

It is assumed that students who were administered unit tests items in sequential order by course objective were able to concentrate their thoughts more closely on the specific objective and would therefore do better on the final course examination than students who were given similar test items in random order by course objective. Further, students who received feedback of 'correct' and 'incorrect' and the correct answer on unit tests would score higher on final examinations and therefore demonstrate the instructional value of the testing mode for computer administered tests. It was assumed that students who received any or different degrees of feedback would therefore score higher on the final exams for the course than students who received only their final percentage score for the unit tests.

In addition to testing the stated hypotheses identified, information was collected regarding the attitudes of students and their preference of the amount and types of feedback provided by each unit test. Information was also collected for the purpose of making comparisons of the students final examination scores in the course and their scores on the ACT test. It was anticipated that students with different levels of ability, as measured by the the composite score on the ACT, would show no variance in benefit from the amount and types of test feedback on computer administered exams.

Definition Of Terms

The following terms are defined as they were used in this study.

SEQUENTIAL ORDER - The presentation of unit test items in order by a specific course objective. All test items for objective 1 were administered before test items for objective 2 were selected. Test items requested by the system were selected for one specific objective and were administered by the computer to the student in random item order for a specific objective.

RANDOM ORDER - The presentation of unit test items to students which were created by the computer in random order by course objective and in random order by item classification within the objectives. The test included items randomly selected from a group of objectives for system presentation.

UNIT TESTS - Study tests were available to students as they proceed through the lessons of the course. Computer Assisted Instruction Lessons existed for the course with parallel test items for each instructional lesson or unit. These unit tests were used to evaluate how well students learned the information contained in the CAI lesson, from outside readings or from class lectures. Most unit tests were from 15 to 25 items long.

FINAL EXAMS - The final examination for the course were administered in two modes of presentation. One class received the test created in random order by objective and item. One class received the test items created in sequential order by objective but in random order by item within each specific objective.

COMPUTER ADMINISTERED EXAMS - The PHOENIX computer managed instruction system was used to create and administer unit and final exams to students. The system had a management mode that allowed test items to be selected in order by sequence or random objective and sequential or random item within a given objective. The presentation of the test items also had feedback options which were used to provide support information to the student after the answer had been entered. At the completion of the unit, the system scored the test and calculated a percentage for the individual test as well as accumulated a score for all combined tests. These values were used to determine the test score differences of students in the study.

Limitations Of The Study

The following limitations exist as they affect the study and the generation of data within it.

- (1) The subjects for this study were limited to students who registered for the OCN 110 Oceanography general education course at UNC during the Fall, Winter and Spring quarters of 1986.
- (2) The subjects for the study were not matched for participation by any other criteria other than their registration for the course in one of the quarters for which the data was collected. Group matching was determined by comparing the composite mean scores on the ACT tests of the students enrolled in each class.

- (3) Test data feedback was limited to the PHOENIX system capacity in the CMI mode of operation.
- (4) This study was limited to the effect of feedback on unit test items used to review the information gained in instruction and was not related to the feedback which might be present in the instructional phase of the CAI lesson on the computer system. This was not an interactive tutorial feedback study but a study to determine the worth of test feedback on unit tests to final exam scores for the course. It was an attempt to determine how much learning took place during testing and not from instruction.

Chapter 2

Selected Literature

Many studies have been done on the effect of feedback on instruction and testing (Arkin 1983, Bender 1984, Hail 1983, Hannafin 1982, Kulhavy 1983 Pifer 1981). None, however, have dealt with the combined effects of various types of test item feedback combined with the modes of test item presentation. This study addressed several types of feedback, including full feedback with remediation and no feedback in conjunction with different modes of question presentation. For example, test items were administered in random order by course objective and in sequential order by course objective. In addition, the combined effects of the type of feedback and modes of presentation on test performance were assessed for different levels of student academic ability as measured by ACT scores.

This summary of related literature is presented in several sections. One section deals with the nature of feedback and its effect on learning. The next section addresses the effect of test feedback on student performance and the anxiety level of students during computer test administration. The final section will review the general psychological basis for using Computer Assisted Learning (CAL) or Computer Assisted Instruction (CAI) in an educational environment. Specifically, an attempt will be made to address the pedagogical history and applications of the computer to learning and the necessity of feedback for learning.

The Nature Of Test Feedback

A review of the research literature on test administration and test

feedback identified several studies pertinent to this research problem. A variety of feedback modes or corrective feedback options have been given to students in a wide variety of courses and grades levels. The net result basically the same in all cases. Item feedback during a computer administered test does improve the performance of students on the test.

Using systemized feedback on tests, Hannafin, (1982), found that providing test feedback in traditional paper presentations created greater test score gains for 6th grade math students when compared to instruction and testing with no feedback. This observation is supported by others (Pifer, 1981, Homa, 1984, Kulhavy, 1985, Arkin, 1984, Schimmel, 1983). It also appears that feedback has its greatest impact on test performance when material is well designed and appropriate for the student population being tested (Kulhavy, 1985).

Shimmel (1983), in a meta-analysis of over 72 studies on test feedback, drew the conclusion that correct/incorrect answer feedback or corrective feedback for high school and college students is more effective than no feedback. Shimmel also discovered that providing the correct answers to students on missed test items contributed to increased test performance. In an extension of Shimmel's work, Arkin (1984) found that corrective feedback on tests appeared to be most effective when students were given two chances to identify the correct answer on an exam. This observation was supported by Tait (1984), for pharmacology students who received immediate feedback on test items including correct answers and explanations of any errors if wrong answers were selected.

Related to the usefulness of feedback is the factor of retention of

information provided on tests. Several studies directly addressed the issue of retention of materials on tests where corrective feedback was provided and the usefulness of the feedback to short and long term memory. Hail (1984), concluded that immediate feedback improved the retention of lesson materials when using self-scoring paper and pencil tests with college students in English. In a class of 5th grade Reading, Peeck (1985), demonstrated that the use of immediate informative feedback showed an increase in the retention of material. Peeck used three modes of feedback; immediate feedback, delayed feedback and no feedback. The end result was that immediate feedback was the most successful in terms of student retention of material for the course/test. Contrary to Peeck's results, Lockhart in a similar study with college students, found that the type of feedback and the quality of feedback did not have a significant effect on retention nor a significant effect on error rate. No other research was found supporting these conclusions, however.

Types Of Feedback And Ability Level

In assessing the effects of types of feedback on test performance of students of different ability, several studies were identified which directly addressed the problem. At the elementary school level, Owens (1981), found that students who had access to information feedback on tests to help them determine how well they were doing, did equally well with either reinforced or nonreinforced feedback on test items. Roberts (1985), discovered that students who were provided explanatory statements about why the answer was wrong demonstrated greater retention than students who were presented the right answer.

Several studies were identified that dealt with the use of test feedback

and the ability level of students. Van Oudenhoven (1983), using 3rd grade spelling students, concluded that both low ability and high ability students increased test performance with test item feedback. The low ability students found the test feedback more beneficial, however. In using the individualized instruction format, test item feedback for both the low and high ability student was shown to be more effective than traditional instruction without test feedback.

At the college level, Homa (1984), classified students in a study as high, medium and low ability. The conclusion of this research indicated that all ability levels of student found feedback superior to no feedback, minimal learning occurring when no feedback was present. This conclusion was supported by Bender (1984) when GPA was used to classify students into two ability groups. Again, results indicated that the low ability student gained more benefit in increased test performance from test feedback than did the high ability student. In addition, the low ability student showed greater improvement in their error analysis capabilities than the high ability student. In this study the high ability student made no significant change in their achievement due to test item feedback. It would appear that the high ability student will get the information and retain it under almost any mode of delivery and feedback system.

Arkin (1983) demonstrated that students in a general education program at the university level do better when they are informed of their level of progress compared to those who do not know how well they are doing. In dealing with the anxiety level of high and low ability students in the testing environment, high ability students do better when they know their progress and test scores whereas low ability students appeared to do better

when they were uninformed. This would lead one to believe that low ability students don't want to know their progress but when given information which provides it, they seem to do better in the course.

Attitudes Toward CMI Test Feedback

The attitudes of students toward testing, CMI test administration and computerized evaluation have been studied rather thoroughly. There are several studies that dealt with the nature of feedback from traditional instruction on paper and pencil tests and compared student attitudes to those of students who received feedback on computer administered tests. Gallagher (1970) found that the computer managed testing (CMI) was found to be superior to traditional teacher administered exams with feedback provided by the teacher. This conclusion was based on the students positive attitude toward CMI and the reduction of test anxiety, even though no significant differences were found in test scores, time to complete the course and time on the system. Lawler (1970), supported this conclusion by confirming a lower anxiety level toward CMI testing. In addition, Lawler found that students who used CMI were more familiar with the objectives for the course than students in traditional instruction and that CMI students actually did better on the final exams and exhibited a higher level of achievement in the course.

Several undergraduate studies were done using CMI test administration with feedback being handled in several ways. Vitulli (1982) let students guess until they found the right answer and provided feedback which was appropriate for each wrong answer. It was found that the feedback of correct and incorrect with support information accounted for the largest

difference in test performance.

Plake (1981) administered tests by computer and varied the sequence of presentation by level of difficulty. The easy test items were presented first with the difficult items last. This system was compared to test administration in the spiral/cyclical mode and a random mode. The results of this study showed that males performed better than females in the easy to difficult presentation. Schroth (1984) extended the test feedback process by varying the time between the answer and the feedback response. A delay of 5 to 20 seconds was provided on the feedback response. It was found that the most useful system of feedback consisted of immediate response rather than delayed response.

Hagerty (1970), allowed students to choose the sequence of their test administration as compared to having the sequence predetermined. It was found that letting students choose the time and type of test administration the performance on midterm and final exams was better than if the teacher had pre-determined it, although the results were not statistically significant. Whetton (1983) allowed students on a verbal ability test in high school to select answers until they found the right answer. The 'answer until correct' (AUC) system enhanced the reliability of the test but did not improve its validity.

A study conducted by Lundeen in 1975, investigated the utility of self-correcting tests in a special education class at UNC. General conclusions reached related to the improved attitude of students toward testing and possible gains in retention of the correct answer. Overall, it was concluded that the advantages of self-scoring paper tests outweigh the disadvantages.

By using the computer to overcome many the of shortcomings of the paper tests, further advantages might be created.

Summary

In summary, the research in the field of test administration and modes of item presentation for computer managed testing offers the following conclusions:

- (1) Corrective feedback does improve the achievement level of students.
- (2) Retention of course materials is enhanced with the use of corrective feedback on test items exists.
- (3) Corrective feedback to students on tests should provide the correct answer and the reasons why any given item was missed.
- (4) Computer generated exams appear to be most beneficial to students when students can determine when to be tested and over what what objectives.
- (5) Students seem to have a positive attitude toward CMI systems when used for test administration.
- (6) Immediate feedback is better than delayed feedback if retention and learning achievement is to be enhanced.
- (7) Low ability students seem to benifit more than high ability students from corrective test feedback.
- (8) Students do better when they know their progress in a course or on a test system than those who receive no feedback.

CHAPTER 3

METHODS AND PROCEDURES

This study was conducted at the University of Northern Colorado during Winter, Spring and Fall quarters of 1986. Students enrolled in the course Oceanography 110 were the subjects of the study. Students enrolled in the course Fall quarter 1986 were selected as the control group for the study. Students enrolled in Oceanography 110 during the Winter and Spring Quarters of 1986 were selected as the experimental groups which would have the feedback limitations and modes and/or of presentation imposed on their unit tests. The experimental and control groups enrolled 696 student in these three sections.

The instructional materials and test items were created by several instructors in Oceanography over the past several years. Test items were validated (content validity) by the instructional cadre in the course and were instructor created in line with the CAI instructional tutorials designed for the course. The test items (files) were created in a parallel structure with the instructional items (files) from the CAI lessons. That is, the test items for Objective 1 of Lesson A covered the same objective as the instructional tutorial for Objective 1 of Lesson A. It should also be noted that the test items were created at the knowledge/recall and comprehension levels of the cognitive domain.

To obtain a base value of achievement scores on unit tests and final exams for the course, all classes were provided tutorial instruction on the lessons for the course without limitations on the number of times any lesson could be taken. Any student could take the CAI lesson as many times as they

wished. The only restriction imposed on the students was the time period that a student could be on the system at any one test presentation. To make opportunities similar for all students, any student enrolled in the class was limited to a 60 minute session as controlled by PHOENIX. If an attempt was made by the student to stay on the system longer than the allotted time, the system would lock-out the student for a 2 hour period. This restriction was imposed on the test files as well. Students were encouraged to take lessons and tests on short 50-60 minute sessions several times rather than one session over a long period of several hours.

The purpose of the study was to determine the effect of feedback regarding performance on unit tests in relationship to final exam scores, as administered by PHOENIX. Thus, the same limitations were imposed on the control and the experimental groups regarding CAI presentation and administration of examinations. Only the unit test feedback and mode of test administration was altered for all the groups.

All students enrolled in the control group were provided the following options in unit test administration. They were given total feedback as provided by PHOENIX each time a test was administered. This total feedback was the 'correct/incorrect' prompt and the correct answer for each item missed. In a sense, students were given only one chance to answer the item correctly and if they missed it, they were told the correct answer. All unit tests administered to students in the control group were administered in random objective order, in random item order with the full feedback.

Exams administered to the experimental groups were as follows:

One group was given the unit exams in sequential order by objective with full

feedback of right or wrong and the correct answer. The other group was given the unit exams in sequential order by objective and random by item with no feedback of right or wrong or the correct answer. Final exams administered to all groups were presented in sequential order by objective with full feedback.

All test items, both unit tests and final tests in the control group, were created from the same test item pool. Each student received a unique test each time a test was administered by PHOENIX. The test pool contained approximately twice the number of items requested for any given test. That means that at worst case senerio, two tests tries for any given unit and two tries on any final exam could provide all of the test items from the pool.

To insure that some students in any one group were not put at a disadvantage by altering the test feedback and final exam scores, every student was given two chances to take the final exam. The highest score was used in the determination of the final letter grade. Where unit test item feedback was withheld, all students in the class were graded in the same manner. Final student grades for the course were adjusted after the experimental differences were identified and recorded.

Because the sample of the study was very large, (three groups for a total of 696 students), it was anticipated that any variation in treatment within any group would be rather easy to identify. Also, because each of the three groups had ACT scores on file required for admission to UNC, it was thought comparisons between the ability level of the groups could be made to see if significant variations in ability existed between the groups. To establish the homogeneity of the population, the composite scores on the ACT test for each of the groups were tabulated and mean values established. A comparison of these mean scores helped determine the similarity of the groups in the

study.

To process the data in the study the SAS statistical package was used. Data obtained in the findings of the study were obtained from means, standard deviations and analysis of variance between dependent and independent variables obtained from the data of the groups.

CHAPTER 4

FINDINGS OF THE STUDY

Introduction

Based on the data obtained from the students participating in the study, the following information is presented. Data are displayed in graph or table form so comparisons can be made between groups. All participants of the study are included in the total figures reported except where noted in the table or graph. Some participants did not have ACT scores available. In computing the relationships between variables where ACT scores were used, only those students who had scores were used in the computation.

Demographic Information

The total number of students in the study was 696. As shown in Table 1, the distributions of males and females between the three groups were essentially the same.

TABLE 1

Population Of The Study

Sex	Group 1 *	Group 2	Group 3	Total
Male	109	131	129	369
Female	91	129	107	327
Total	190	260	246	696

- * 1) Group one received random objective, random item, full feedback
 2) Group two received sequential objective, random item, full feedback
 3) Group three received sequential objective, random item, no feedback

Data in Table 2 identifies the class standing of the students enrolled in the three Oceanography 110 classes Winter 1986, Fall 1986 and Spring 1986.

Table 2
Class Makeup Of The Population In The Study

Class	Number	Percent
Freshman	369	54.3
Sophomore	185	27.2
Junior	62	9.1
Senior	64	9.4
Unclassified	* 16	-
Totals	696	100.0

* Not included in the classification by grade level but included in the study.

Table 3 presents the numbers of students in the treatment groups who had ACT scores available. Because it was not possible to split each class into control and experimental groups due to the management controls of PHOENIX, entire classes became treatment groups. To determine the compatibility of the groups, ACT scores were used in an analysis of covariance to account for any initial group ability differences. As Table 3 indicates, the group mean ACT scores were very similar. Analysis of variance showed no significant difference between mean composite ACT scores among the groups ($F=.23, p>.05$).

Table 3
Comparison Of ACT Scores Of The Population

	Group 1	Group 2	Group 3
Total Population	190	260	246
ACT Scores Available	100	146	120
ACT Mean Score	18.04	18.58	18.56
ACT Std. Deviation Value	4.4	4.2	4.5
Minimum ACT Score	9.0	9.0	7.0
Maximum ACT Score	28.0	29.0	27.0
ACT English Mean Score	17.17	17.98	18.11
ACT Math Mean Score	17.02	17.25	16.93
ACT Soc. Stud. Mean Score	16.51	17.40	17.31
ACT Science Mean Score	20.97	21.09	21.40

To make further comparisons between the treatment groups, the mean final grades and final exam scores are presented in Table 4. Notice that the 'Final' represents the mean score on the PHOENIX administered final examination for the course for each group. The 'Grade' is the mean class grade calculated for the students in each group. s

Table 4
Final Test Scores And Class Grades Of Groups

	Group 1	Group 2	Group 3
PHOENIX Final Exam Mean	81.48	82.83	80.83
Final Class Grade Mean	3.24	3.02	2.95

Findings Of The Study

Analyses of variance and covariance were used to determine the variations in the test scores due to treatment differences. All calculations were made

with the level of significance pre-set at .05.

Effects Of Unit Feedback On Final Exam Scores

To determine if final exam scores had any relationship to the type and amount of feedback given to CAI students on unit tests administered by PHOENIX, several variations of test feedback on unit tests were set up. It was assumed that feedback in a practice testing environment would contribute to the level of performance of students receiving such feedback. To examine the specific effects of feedback, several variations were created on the PHOENIX delivered practice unit tests. Group 1 received full feedback including the correct/incorrect response and the correct answer if an item was missed on the unit tests. Students were given only one chance to answer the item correctly. The entire course had 21 unit tests which contained 10 items each. If a student took each test only once, a minimum of 210 test feedback frames were presented. On the average, students took the tests two times which increased the number of feedback items from 210 to 420. Each test took about 8 minutes to complete so extensive computer time was not involved with the testing. Students could take the tests as many times as they wished. The test data file contained over twice as many items as were presented in any one test. Each time a student took the test, the items on the tests were different as the test generator in PHOENIX presented the test items in random order from the pool of questions.

In comparing all three groups in regards to feedback and mode of presentation (sequential vs. random objective), the greatest gain in final exam score was achieved when sequential presentations by objective and full feedback was used. An analysis of co-variance indicated that the feedback

was more valuable than the mode of presentation although both were helpful. Table 5 presents the adjusted means in the analysis of co-variance for the three groups with the final score being the dependent variable. These results indicate that the feedback on individual test items was more helpful than no feedback and that sequential objective presentations were more helpful than random objective presentations.

With significant differences found among the adjusted means, pair wise controls were examined using Sheffe comparisons. As indicated in Table 4, Group 2 students who received their test questions in sequential order by objective and in random order by item, and students who received full feedback of correct/incorrect for any missed item, received the highest scores on their final exams. Group 1 received their unit tests in random order by objective and in random order by item with full feedback of correct/incorrect. The difference in final exam scores for those two groups, although not significant, was 81.48% for random presentation vs. 82.83% for sequential presentation. Presentation of test items in sequential objective order is better than presenting test items in random objective order.

The difference between the groups who received their test items in sequential order but with the feedback as the experimental variable was also found to be significantly different. Group 2 received full feedback of test items (correct/incorrect with the correct answer given if wrong) and Group 3 received no feedback whatsoever during the test presentation. At the end of the test, Group 3 were given their scores for the test in percentage form as the only record of how they did on the test. The difference in mean scores between Group 2 (full feedback) at 82.83% and Group 3 (no feedback) at 80.83%

was 2% points. This difference was determined to be significant at the .05 level.

Table 5
Analysis Of Covariance
Final Exam Scores For Three Treatment Groups

Source	df	Adjusted SS	F
Treatment Groups	2	516.11	3.72 *
ACT Covariate	1	2885.71	41.58 *
Error	243	16864.56	

Significant at .05

Based on an analysis of data on the effects of feedback on unit test items, full feedback is more productive than no feedback. This conclusion, coupled with the presentation of items in sequential order by objective, demonstrates that students will gain more information if the test items are presented in sequential order by objective with full feedback provided. This feedback is necessary if the answer is right or wrong. Feedback which explains the error if wrong and supports the correct choice with additional data if right, provides for the greatest amount of gain in the learning experience. Those modes which simply identify the answer as right or wrong showed better scores than the modes of no feedback. Adding supported information to the answer extended the effectiveness of the feedback and provided for higher final exam scores than just identification of the correct answer.

Effects Of Unit Feedback On Final Exam Scores

To determine if final exam scores had any relationship to the type and amount of feedback given to CAI students on unit tests administered by PHOENIX, several variations of test feedback on unit tests were set up. It was assumed that feedback in a practice testing environment would contribute to the level of performance of students receiving such feedback. To examine the specific effects of feedback, several variations were created on the PHOENIX delivered practice unit tests. Group 1 received full feedback including the correct/incorrect response and the correct answer if an item was missed on the unit tests. Students were given only one chance to answer the item correctly. The entire course had 21 units tests of 10 items each. If a student took each test only once, a minimum of 210 test feedback frames were presented. On the average, students took the tests two times which increased the number of feedback items from 210 to 420. Each test took about 8 minutes. Students could take the tests as many times as they wished. Each test data file contained more than twice as many items as were presented in any one test. Each time a student took a test, that particular test was unique as the test generator in PHOENIX presented the test items in random order from the pool of questions.

Mode Of Presentation

In comparing the three groups relative to feedback and mode of presentation (sequential vs. random objective), final exam scores were highest when sequential presentations by objective and full feedback was used. An analysis of co-variance indicated that the feedback was more valuable than the mode of presentation although both were helpful. The adjusted means in the

analysis of co-variance for the three groups with the final score being the dependent variable indicated that the feedback on individual test items was more helpful than no feedback and that sequential objective presentations were more helpful than random objective presentations.

Pair wise differences among the adjusted means were examined using Sheffe comparisons. Students who received their test questions in sequential order by objective and in random order by item, and students who received full feedback of correct/incorrect for any missed item, received the highest scores on their final exams. The difference in final exam scores for the two groups, although not significant, was 81.48% for random presentation vs. 82.83% for sequential presentation.

The makeup of the three groups in the study allowed for comparisons between different types of feedback and modes of presentation. One comparison was made between groups who had the same mode of presentation (sequentially by objective) but different types of feedback (full vs. none). Another assessment was made by comparing groups who were given the same type of feedback (full feedback) but were presented the test items in different modes (sequential order by objective vs. random objective order).

To make the comparisons between the effect of feedback to final exam scores, the mode of presentation was held constant for groups 2 and 3. Both groups were presented test items in sequential order by objective and random order within objective. Group 2 received full feedback of test items (correct/incorrect with the correct answer given if wrong) and Group 3 received no feedback whatsoever during the test presentation. Delayed feedback was provided to Group 3 at the end of the test in percentage form.

CHAPTER 5

Summary Of The Study, Findings, Conclusion Implications And Recommendations

The purpose of this study was to compare the academic performance of college students responding to combinations of test item feedback methods and modes of test item presentation to final examination scores. All examinations and feedback were administered by the PHOENIX computer based education system. More specifically, the study attempted to answer the following questions.

- (1) Will students who receive partial feedback of 'correct' or 'incorrect' and are provided the correct answer on unit test items score higher on final examinations than students who receive no feedback?
- (2) Will students who are presented unit test items in sequential order by course objective score higher on final examinations than students who are presented the unit test items in random course objective order?
- (3) What relationship exists between the types and amount of feedback on unit test items presented in random order by course objective to student ACT scores and final examination scores for the course?
- (4) What relationship exists between the types and amount of feedback on unit test items presented in sequential order by course objective to the ACT scores and final examination scores of students in the course?

Review of Literature Nature Of Test Feedback

A review of the research literature on test administration and test feedback identified several studies pertinent to this research problem. A variety of feedback modes or corrective feedback options have been given to students in a wide variety of courses and grade levels. The net result has been found to be basically the same in all cases. That is, item feedback

during a computer administered test does improve the performance of students on the test.

A summary of research in the field of test administration and modes of test item presentation for computer managed testing offers the following conclusions:

- (1) Corrective feedback does improve the achievement level of students. (Arkin, 1983), (Bender, 1984), (Hail, 1983)
- (2) Retention of course materials is enhanced with the use of corrective feedback on test items. (Hail, 1983), (Peek, 1985)
- (3) Corrective feedback to students on tests should provide the correct answer and the reasons why any given item was missed. (Roberts, 1985), (Jelden, 1986), (Vitulli, 1983)
- (4) Computer generated exams appear to be most beneficial to students when students can determine when to be tested and over what objectives. (Arkin, 1983), (Hagerty, 1970)
- (5) Students seem to have a positive attitude toward CMI systems when used for test administration. (Gallagher, 1970), (Lawler, 1970)
- (6) Immediate feedback is better than delayed feedback if retention and learning achievement is to be enhanced. (Schroth, 1984), (Plake, 1981)
- (7) Low ability students seem to benefit more than high ability students from corrective test feedback. (Homa, 1984), (Bender, 1984)
- (8) Students do better in course achievement when they know the status of their progress in a course or on a test than those who receive no feedback or course progress information. (Arkin, 1983)

Methods And Procedures

The data for the study were obtained from 696 students enrolled at UNC in the Oceanography 110 course Fall, Winter and Spring quarters 1986. Fall quarter 1986 was selected as the control group for the study. Students enrolled in the course during the Winter and Spring quarters of 1986 were selected as the experimental groups which would have the feedback limitations

and/or modes of presentation imposed on their unit tests.

The instructional materials and test items were created by several instructors in Oceanography. Test items were validated (content validity) by the instructional cadre in the course and were instructor created in line with the Computer Assisted Instruction (CAI) interactive tutorials designed for the course. The test items (files) were created in a parallel structure with the instructional items (files) from the CAI lessons. That is, the test items for Objective 1 of Lesson A covered the same objective as the instructional tutorial for Objective 1 of Lesson A. It should also be noted that the test items were created at the knowledge/recall and comprehension levels of the cognitive domain.

To obtain a base value of achievement scores on unit tests and final exams for the course, all classes were provided tutorial instruction on the lessons for the course without limitations on the number of times any lesson could be taken. Any student could take the CAI lesson as many times as they wished. The only restriction imposed on the students was the time period that a student could be on the system at any one test presentation. To make opportunities similar for all students, any student enrolled in the class was limited to a 60 minute session as controlled by PHOENIX. If an attempt was made by the student to stay on the system longer than the allotted time of 60 minutes, the system would force the student out of the lesson and lock them out for a 2 hour period. This restriction was imposed on test files as well. Students were encouraged to take all lessons and tests in short 50-60 minute sessions rather than a few extended sessions of several hours.

The purpose of this study was to determine the effect of feedback regarding performance on unit tests in relationship to final exam scores, as

administered by PHOENIX. Thus, the same limitations were imposed on the control and the experimental groups regarding CAI presentations and administration of examinations. Only the unit test feedback and mode of test administration was altered for all groups in the experiment.

All students enrolled in the control group were provided the following options in unit test administration. They were given a total feedback as provided by PHOENIX each time a test was administered. This total feedback was the 'correct/incorrect' prompt and the correct answer for each item. Students were given only one chance to answer the item correctly and if they missed it, they were told the correct answer. All unit tests administered to students in the control group were administered in random course objective order, in random item order within the objective and provided with full item feedback.

Examinations administered to the experimental groups were as follows: One group was given the unit exams in sequential order by course objective with full feedback of right or wrong with the correct answer. The other group was given their unit examinations in sequential order by course objective and random order by item with no feedback of right or wrong or the correct answer. Final examinations administered to all groups were presented in sequential order by course objective, random order by item within the objective and were displayed with full feedback.

All test items, both unit tests and final tests, were created from the same test item pool. Each student received a unique test each time a test was administered by PHOENIX. The test pool contained approximately twice the number of items requested for any given test. That means that at worst case

senerio, two test tries for any given unit and two tries on any final exam could provide all of the test items from the pool.

To insure that some students in any one group were not put at a disadvantage by altering the test feedback and final exam scoring system, every student was given two chances to take the final exam. The highest score obtained on either test was used in the determination of the final letter grade for the test. Where unit test item feedback was withheld, all students in the classes were graded in the same manner. Final student grades for the course were adjusted after the experimental differences were identified and recorded.

Because the sample of the study was very large, (three groups for a total of 696 students), it was anticipated that any variation in treatment within any group would be rather easy to identify. Also, because each of the three groups had ACT scores on file required for admission to UNC, it was thought comparisons between the ability level of the groups could be made to see if significant variations in ability existed between the groups. To establish the homogeneity of the population, the composite scores on the ACT test for each of the groups were tabulated and mean values established. A comparison of these mean scores helped determine the similarity of the groups in the study.

To process the data in the study, the SAS statistical package was used. Data obtained in the findings of the study wee obtained from means, standard deviations and analysis of variance between dependent and independent variables obtained from the data of the groups.

Findings Of The Study

Demographic Information

The total number of students in the study was 696. The numbers of males and females in each of the three groups were essentially equal. Half of the participants were freshman and about one-fourth were sophomores. The other one-fourth were equally distributed between juniors and seniors. Because it was not possible to split each class into control and experimental groups due to the management controls of PHOENIX, entire classes became treatment groups. To determine the compatibility of the groups, ACT scores were used in an analysis of covariance to account for any initial group ability differences. The group mean ACT score were very similar. Analysis of variance showed no significant difference between mean composite ACT scored among the groups ($f=.23$, $p>.05$).

To make further comparisons between the treatment groups, the mean final grades and final exams scores were compared. Analysis of variance and covariance were used to determine the variations in the test scores due to treatment differences. All calculations were made with the level of significance pre-set at .05.

Effects Of Unit Feedback On Final Exam Scores

To determine if final exam scores had any relationship to the type and amount of feedback given to CAI students on unit tests administered by PHOENIX, several variations of test feedback on unit test administration were set up. It was assumed that feedback in a practice testing environment would contribute to the level of performance of students receiving such feedback. To examine the specific effects of feedback, several variations were created

on the PHOENIX delivered practice tests. Group 1 received full feedback including the correct/incorrect response and the correct answer if the item was missed on the unit test. Students were given only one chance to answer the item correctly. The entire course had 21 unit tests of at least 10 items each. If a student took each test only once, a minimum of 210 test feedback frames were presented. On the average, students took the test two times which increased the number of feedback items from 210 to 420. Each test took about 8 minutes. Students could take each unit test as many times as they wished. All test items were generated from the test item pool created for that unit.

The first major hypothesis to be tested with the study was:

Hol - Students who are presented unit test items in sequential order by course objective will score significantly higher on final exams than students who are presented the unit test items in random order by course objective.

Based on an analysis of data obtained in the study, this was found to be true. Differences in terms of point values were not great (2 percentage points), but it was found to be significant at the .05 level.

The second hypothesis of the study related to the degree of test feedback and its usefulness on final exam scores. The hypothesis stated was:

Ho2 - Students who are given feedback on unit tests items of 'correct' and 'incorrect' as well as the right answer will score significantly higher on final examinations than students who are provided no feedback on unit tests.

Data obtained in the study held this hypothesis to be true. Again, the differences were not great in actual numbers (2 percentage points), but it was statistically significant at the .05 level.

Mode Of Presentation

In comparing the three groups relative to feedback and mode of

presentation (sequential vs. random objective), final exam scores were highest when sequential presentations by course objective and full feedback were used. An analysis of co-variance for the three groups with the final score being the dependent variable indicated that the feedback on individual test items was more helpful than no feedback and that sequential objective presentations were more helpful than random objective presentations.

Pair wise differences among the adjusted means were examined using Sheffe comparisons. Students who received their test questions in sequential order by objective and in random order by item, and students who received full feedback of correct/incorrect for all items, received the highest scores on their final exams. The difference in final exam scores for the two groups, although not great, was 81.48% for random presentation vs. 82.83% for sequential presentation. This difference was significant at the .05 level.

The makeup of the three groups in the study allowed for comparisons between different types of feedback and modes of presentation. One comparison was made between groups who had the same mode of presentation (sequentially by objective) but different types of feedback (full vs. none). Another assessment was made by comparing groups who were given the same type of feedback (full feedback) but were presented the test items in different modes (sequential order by objective vs. random objective order).

To make the comparisons between the effect of feedback to final exam scores, the mode of presentation was held constant for groups 2 and 3. Both groups were presented test items in sequential order by objective and random order within objective. Group 2 received full feedback of test items (correct/incorrect with correct answer given if wrong) and Group 3 received no feedback whatsoever during the test presentation. Delayed feedback was

provided to Group 3 at the end of the test in percentage form. This was the only means of feedback provided relative to test performance. The difference in mean scores between Group 2 (full feedback) at 82.83% and Group 3 (no feedback) at 80.83% was 2% points. This difference was determined to be significant at the .05 level.

To make comparisons between the modes of presentation of test items, with the type and amount of feedback held constant, test items for Group 1 were generated in random order by course objective and item. This means that the tests were fully random creations, both by objective and by item within each objective. Group 2 had test items generated in sequential order by objective but in random order within objective (5 test items for each objective, for example, were presented to each student in random order within that objective). By comparing the performance of students in these two groups, it was found that the sequential objective testing of Group 2 was superior to the random objective testing of Group 1. Although the mean scores for each group (Group 2 = 82.83% and Group 1 = 81.48%) was not great, the difference was significant at the .05 level. In comparing the differences between both modes and types/amounts of feedback, it was found that the feedback difference was greater than that created by the mode of presentation. The degree of feedback has more impact on the final exam scores of students than the mode of presentation of test items. By combining the sequential objective order of test presentation with full remedial feedback for each test item presented, the greatest gain on final test scores was achieved.

To summarize the general results of the finding of this section, the statement of hypothesis tested was:

Ho3 - Students who are administered unit test items in sequence by objective and given feedback of 'correct' and 'incorrect' as well as the right answer will score significantly higher on final exams than students given similar information on unit test items which were administered in random course objective order.

Data obtained from the study did prove this hypothesis to be true. As with the previous findings, the differences were numerically small but they were significant at the .05 level.

The last hypothesis to be tested dealt specifically with the mode of presentation. It was stated as follows:

Ho4 - Students who are administered unit test items in sequential order by course objective and provided no item feedback will score significantly higher on final exams than students who are administered similar tests items in random order by course objective.

The data for the study again supported the hypothesis. Although the differences were much less than those related to feedback, they were still significant at the .05 level. It would appear that feedback is more important than modes of presentation to achievement on final examination scores.

Conclusions, Implications And Recommendations

The following conclusions, implications and recommendations are presented based on the finding of this study and from a synthesis of related studies identified as supporting the problem of test item feedback and modes of item presentation.

- (1) Corrective feedback does slightly improve the achievement scores of students as measured by objective tests administered by computer.

Teachers who use a computer to construct sample or practice tests as well as final exams for students should use feedback for both correct and incorrect answers as a means for supporting instruction. When students are learning information for course work, corrective feedback and problem explanations

for incorrect answers are effective ways to extend an understanding of the concepts under study.

The finding of this study and of others analyzed imply that the more feedback provided to students in testing, that is providing the rationale for why answers are wrong or right, the better the learning and the longer it will be remembered. Retention of course materials is enhanced with the use of corrective feedback on test items. This would imply that students who receive corrective feedback on sample or instructional tests learn the material better and remember it longer. When teachers develop instructional materials, they should attempt to use this type of presentation with all evaluation items. On the other hand, test feedback may not be appropriate for initial competency testing of students. If teachers use competency testing to establish instructional procedures, partial feedback of right or wrong be be appropriate to establish the learning set for students.

- (2) Computer generated examinations appear to be most beneficial to students when they have control over the speed of the item presentation and the objectives over which they are to be tested.

Informal discussions with students in this study and an analysis of previous research which deal with attitudes of students toward testing show that attitudes of students toward computer administered examinations are better when students feel they have control over when tests are given and have some input over what objectives are to be examined. CMI system developers should consider providing these options to students when developing software authoring packages. The general attitude of students toward testing is improved when students feel some control over the timing, mode and management features of the testing structure. Teachers should establish what to teach

and what to test and students should have input to when and how testing is done.

- (3) Immediate feedback on specific test items is more beneficial to students than delayed feedback. Retention of materials is enhanced with immediate feedback.

Instructors who administer tests on the computer should provide immediate feedback to students on every test item. Any CMI system generating test items should have provision for at least correct/incorrect feedback for answers on tests as the test is administered. Remedial feedback provision is even better support and improves the learning process.

- (4) Low ability students seem to benefit more from corrective test feedback than do high ability students.

This research study looked at ability groupings of students and final test scores for the various modes and test feedback presentations. No significant difference was found by ability groupings when identified by ACT scores and final examination scores. Previous research studies, however, indicate that low ability students appear to gain more from test feedback than high ability students. This would imply that low ability students gain more benefit from corrective feedback than high ability students. Although the differences did appear in this study, they were not statistically significant. This would imply that more research needs to be done specifically related to ability levels of learners and feedback for computer administered examinations. Related to this problem is the modes of presentation as it relates to test administration and delivery by computer. The findings of this research, which states that all students do better on a test when item feedback is provided on practice unit tests, implies that teachers and instructional developers should implement a progress record as well as a

feedback structure within any computer administered test system. Knowledge of results is important to students as learning takes place but whether computer administered tests are equally beneficial to all students or more important to low ability students has yet to be specifically validated.

- (5) Tests generated by sequential course objective and in random order by item within that objective yield higher scores than tests created by random objective order.

This conclusion implies that teachers should present information to students in a planned format generally by objective. If instruction follows a logical sequence which is psychologically ordered, then testing should follow that same structure. Testing to be effective and efficient, should follow sequential objective order and will produce higher test scores on tests administered by computer. If students are presented items on a test over a specific objective, they have a tendency to relate the information better than when test items are presented in random objective order. Developers of CMI test generation systems should take into consideration the ability of asking the systems to generate tests in sequential/random order by course objective as well sequential/random order by test item. With this configuration, the teacher will have the option of using test presentation mode in either format. The general conclusion of this study is that students perform better on final objective tests if they have instructional tests with corrective feedback presented in sequential order by course objective preceding the formal final exam. Tests with items clustered by objective result in higher final exam scores than tests presented in random objective and random item order.

BIBLIOGRAPHY

- Arkin, Robert M. and Schumann, David W., "Effect Of Corrective Testing: An Extension", (University of Missouri - Columbia). Journal of Educational Psychology, October 1984, Volume 76 (5), pp. 835-843.
- Arkin, Robert M. and Qalts, Elizabeth A., "Performance Implications Of Corrective Testing", (University of Missouri -Columbia). Journal of Educational Psychology, August 1983, Volume 75(4), pp. 561-571.
- Bell-Gredler, Margret E., "Learning And Instruction; Theroy Into Practice", Macmillan Publishing Company, New York, 1986.
- Bender, Timothy A., "Improving The Use Of Classroom Feedback", Presentation American Educational Research Association, 68th Conference, New Orleans, LA., April 23-27, 1984. Conference Proceedings.
- Brown, Frederick G., "Principles of Educational And Psychological Testing", Holt, Rinehart and Winston, New York, 1983.
- Gallagher, Paul D., "An Investigation Of Instructional Treatments And Learner Characteristics In Computer-Managed Instruction Course", Florida State University, CAI Center, Tallahassee, Florida. Technical Report No. 12, July 30, 1970, pp. 94.
- Hagerty, Nancy K., "Development And Implementation Of A Computer-Managed instruction System In Graduate Training", Florida State University, CAI Center, Tallahassee, Florida. Technical Report No. 11, June 30, 1970, pp. 161.
- Compress Quarterly, "A Compilation Of Views On Educational Software", Volume 2, Number 4, Fall 1985.
- Hail, Jimmy. G., "The Effects Of Using Immediate Feedback Answer-Until-Correct Test Procedures On Retention", (Baylor University - Waco, Texas). Dissertation Abstracts International, January 1983, Volume 43 (7-A), #2285.
- Hannafin, Michael J., "The Effects Of Systemized Feedback On Learning In Natural Classroom Settings", (University of Colorado -Boulder). Educational Research Quarterly, Fall 1982, Volume 7(3), pp. 22-29.
- Homa, Donald and Cultice, Joan, "Role Of Feedback, Catagory Size, and Stimulus Distortion On The Acquistion And Utilization Of Ill Defined Categories", (Arizona State University -Tempe). Journal of Experimental Psychology: Learning, Memory, & Cognition, January 1984, Volume 10 (1), pp. 83-94.
- Jelden, David L., "Effectiveness Of Computer Based Education: The UNC PHOENIX System". (University Of Northern Colorado - Greeley, CO.)

Test Feedback Study - Jelden

Staff Study - 1986 Available From ERIC (EDRS), 118 pgs.

- Kulhavy, Raymond W., "Feedback Complexity And Corrective Efficiency". Contemporary Educational Psychology, Vol. 10(3), July 1983, pp. 285-291.
- Lawler, R. Michael, "An Investigation Of Selected Instructional Strategies in An Undergraduate Computer-Managed Instruction Course". Florida State University, CAI Center, Tallahassee, Florida. Technical Report No. 19, April 30, 1971. pp. 107.
- Lockhart, Kathleen A., et.al., "The Effects If The Timing Of Feedback On Long Term Knowledge Retention In PSI Courses. Final Report". Navy Personnel Research and Development Center, San Diego, CA. Report #NPRDC-TR-83-13, US GPO, February 1983, 34 pgs.
- Lundeen, Dale J., "An Investigation Of The Utility Of A Self-Scoring Test", Colorado Journal Of Educational Research, Volume 15, pp. 2-4, 1975.
- Oosterhof, Albert C. and Salisbury, David. F., "Some Measurement and Instruction Related Considerations Regarding Computer Assisted Testing", Presentation, American Educational Research Association, 68th Annual Convention, New Orleans, LA., April 23-27, 1984. Convention Proceedings pp. 14.
- Owens, Carl R., et.al., "The Function Of Direct And Vicarious Reinforcement In Human Learning". Conference Presentation, Southwestern Psychological Association, 27th Meeting - April 16-18, 1981, Houston, TX. Available From ERDS-12 pgs.
- Peeck, J. et.al., "Effects Of Informative Feedback In Relation To Retention Of Initial Responses", Contemporary Educational Psychology, October 1985, Volume 10(4), pp. 303-313.
- Pifer, Raymond E., "Effects Of The Use Of Feedback On Achievement", (University of Chicago). Dissertation Abstracts International, September 1981, Volume 42(3A), #916.
- Plake, Barbara S.; Melican, Gerald J.; Carter, Linda; Shaughnessy, Michael J., "Differential Performance Of Males And Females On Easy To Hard Item Arrangements: Influence Of Feedback At The Item Level", (University Of Nebraska - Lincoln). Educational & Psychological Measurement, Winter 1983, Volume 43(4), pp. 1067-1075.
- Roberts, Franklin C. and Park, Ok-Choon, "Feedback Strategies And Cognitive Style In Computer-Based Instruction", (Control Data Corporation - Minneapolis). Journal of Instructional Psychology, June 1984, Volume 11(2), pp. 63-74.
- Schummel, Barry J., "A Meta-Analysis Of Feedback To Learners In Computerized And Programmed Instruction", Presentation - American Educational Research

Test Feedback Study - Jelden

Association, Montreal, Canada - April 11-14, 1983, 17 pgs.

Schroth, Marvin L., "Directive Concept Formation: S-R Hypothesis Versus Cognitive Theory", (University of Santa Clara -California).
Journal of General Psychology, January 1984, Volume 110(1), pp. 87-92.

Tait, K. and Hughes, I.E., "Some Experiences In Using A Computer-Based Learning System As An Aid To Self-Teaching And Self Assessment.
Computers And Education, Volume 8(3), 1984, pp. 271-278.

Van Oudenhoven, Jan P.; Siero, Frans; Veen, Perte; Siero, Sjef, "Effects Of Positive Feedback On Self-Evaluations And Spelling Achievement In Real Educational Setting", (State University - Groningen, Netherlands).
European Journal of Social Psychology, July-September 1982, Volume 12(3), pp. 321-325.

Vitulli, William F., "Effects Of Immediate Feedback On Computer-Assisted Testing Of ESP Performance", (University of South Alabama - Mobile).
Psychological Reports, October 1982, Volume 51(2), pp. 403-408.